УДК 004.934

В.В. Яценко, М.М. Сажок
Міжнародний науково-навчальний центр інформаційних технологій та систем,
м. Київ, Україна
Україна, 03680, просп. Акад. Глушкова, 40, МСП, м. Київ, { yatsenko.valya , sazhok }@gmail.com

Автоматизовані засоби формування лінгвістичної бази даних і знань для системи усного перекладу

V.V. Yatsenko, M.M. Sazhok
Speech Science and Technology Department, International Research and Training Center of Information Technologies and Systems, Kyiv, Ukraine
Ukraine, 03680, Acad. Glushkov Ave., 40, MSP, Kyiv, {yatsenko.valya , sazhok }@gmail.com

Automated Means for Formation of Linguistic Data and Knowledge Base for a Spoken Translation System

В.В. Яценко, Н.Н. Сажок
Международный научно-учебный центр информационных технологий и систем,
г. Киев, Украина
Украина, 03680, пр. Акад. Глушкова, 40, МСП, г. Киев

Автоматизированные средства формирования лингвистической базы данных и знаний для системы устного перевода

У статті розглядаються засоби формування бази даних і знань для системи усного перекладу на основі моделі розпізнавання та інтерпретації мовленнєвого сигналу в межах предметних областей з урахуванням метаслів. Описується структура та програмні модулі інструментарію, який дає змогу оперувати предметними областями, мовами, типами смислів і речень та метасловами. На основі створеної БД і З для трьох мов розроблено демонстраційну систему усного фразника-перекладача.
Ключові слова: розпізнавання та інтерпретація мовлення, фразник-перекладач, предметна область, тип речення, база даних і знань.

This paper describes the formation of data and knowledge base for spoken interpreter system based on the model of speech signal understanding within subject areas including meta-words. The proposed structure allows for operating the subject areas, languages, sentence and meaning types and meta-words. The respective toolkit is developed and described. The experimental system of a spoken phrasebook-interpreter exposed results proper for its practical application. The presented demo-version is provided with the data and knowledge base for a new language created by the developed means.
Key Words: speech recognition and understanding, spoken phrasebook, subject area, sentence type, data and knowledge base.

В статье рассматриваются средства формирования базы данных и знаний для системы устного перевода на основе модели распознавания и интерпретации речевого сигнала в пределах предметных областей с учетом метаслов. Описывается структура и программные модули инструментария, который позволяет оперировать предметными областями, языками, типами смыслов и предложений и метасловами. На основе созданной БД и З для трех языков разработана демонстрационная система устного фразника-переводчика.
Ключевые слова: распознавание и интерпретация речи, фразник-переводчик, предметная область, тип предложения, база данных и знаний.

Вступ

У повсякденному житті людині доводиться розв'язувати досить багато задач, які пов'язані з великим обсягом інформації. Далеко не всі задачі є алгоритмічними. Розв'язок багатьох задач зводиться до керування потоками інформації, аналізу даних. Будь-які знання – це свого роду дані, якими володіє людина. Якщо для розв'язку задач необхідні знання про однотипові об'єкти чи явища, що повторюються, то варто використовувати базу даних (БД). База даних – це структуровані знання про об'єкти.

База даних допомагає систематизувати і зберігати інформацію з певної предметної області, полегшує доступ до даних, пошук та надання необхідних відомостей. Найпростішою БД можна вважати телефонний довідник або список книг у домашній бібліотеці. Сучасні БД оперують інформацією, представленою в розмаїтому форматі – від звичайних чисел і тексту до графічних і відеоданих.

Останнім часом досить популярним стало використання on-line перекладачів, які використовують БД великих обсягів. Компанія Google представила свій веб-сервіс – Google Перекладач (англ. Google Translate), призначений для перекладу тексту чи веб-сторінки на іншу мову [1]. Цей сервіс, як і інші інструменти автоматичного перекладу, має свої обмеження і недоліки. Постійно ведеться вдосконалення якості перекладу, оскільки цей інструмент дає можливість користувачеві зрозуміти загальний зміст тексту, він не дає точний і граматично правильний переклад; складається враження, що переклад виконується без урахування семантики і прагматики.

Використання комп'ютерів для перекладу переважно ґрунтується на роботі із словниками та глосаріями. Досить популярним в наш час стало використання пам'яті перекладів (англ. Translation Memory, TM), що містять приклад раніше перекладеного тексту термінологічними базами, а також використання текстових корпусів – великих колекцій текстів однією чи багатьма мовами, що дає стислий опис того, як слова та вирази реально використовуються в мові в цілому чи в конкретній предметній області.

При побудові систем усного перекладу в межах предметних областей доводиться використовувати та опрацьовувати досить великі обсяги інформації. Зв'язування та структурування представленої інформації формує базу даних, з якою може працювати експерт-лінгвіст. База даних є інформаційною моделлю предметної області.

У даній роботі ведеться мова про формування, поповнення та збереження не просто бази даних, а бази даних і знань (БД і З). У компоненті БД і З, що стосується знань, описуються різноманітні правила і закономірності взаємозв'язку тексту і звуків (фонем), семантики та лексикону та ін.

При формуванні системи усного перекладу виникає ряд проблем, пов'язаних із задачами розпізнавання та розуміння мовленнєвого сигналу, а саме, необхідно вміти будувати моделі всіх можливих речень мови діалогу, що виражають один і той самий зміст, вміти моделювати параметри слів у типах речень, генерувати та знаходити найбільш правдоподібні еталонні сигнали, враховуючи параметри. Це дає змогу сформувати для розпізнаної фрази якомога точний відповідник іншою мовою навіть у випадку, якщо відповідь розпізнавання містить помилки. Вирішенню цих питань були присвячені попередні статті [2], [3].

Тепер пропонується розглянути питання автоматизації формування даних і знань, необхідних для системи усного перекладу. В рамках запропонованої структури передбачається можливість вводити та обирати вхідну мову для інтерпретації та вихідну мову для перекладу, редагувати, створювати та поповнювати предметні області (ПО). З огляду на це подається опис функціональної схеми, яка призначена для автоматизованого формування БД і З для усного фразника-перекладача.

Предметні області виконують ті ж функції, як і у паперовому розмовнику. Кожна ПО містить моделі речень, що є різними за формою, але виражають один і той же зміст. Передбачається можливість працювати з уже існуючими ПО та поповнювати їх новими даними. Таким чином, створюється інструментарій, за допомогою якого експерт-лінгвіст формує дані предметних областей для своєї мови, тим самим забезпечуючи функціонування перекладу для всіх пар мов, що включають цю мову.

Спочатку розглянуто загальну структуру системи усного перекладу, з якої випливає структура БД і З, потім описується інструментарій для формування БД і З, насамкінець подаються експериментальні дослідження, що супроводжуються висновками і планами на майбутнє.

Структура системи усного перекладу в межах предметних областей

Система усного перекладу реалізована на основі задач розпізнавання та змістовної інтерпретації злитого усного мовлення. Ці задачі виконуються в єдиному взаємопов’язаному процесі, кінцевою метою якого є зміст повідомлення, який передається послідовністю слів, та його переклад іншою мовою.

Суть задачі розпізнавання мовленнєвого сигналу полягає в тому, що це є процес автоматичного аналізу сигналу з метою визначення послідовності слів, які передаються цим сигналом. А задача змістовної інтерпретації мови – це процес автоматичного оброблення мовленнєвого сигналу з метою виявлення змісту, що передається сигналом, та представлення цього змісту в певній канонічній формі, зручній для подальшого використання в системі усного перекладу [2-4].

Для системи усного перекладу задача змістовної інтерпретації є ключовою. Оскільки для множини речень, які можуть передавати одну й ту саму думку, треба відтворити однозначний зміст, за яким потім отримати переклад на іншу мову. У попередній роботі [3] під поняттям «параметр» ми розуміли опис множини можливих варіантів висловів у фразах при спілкуванні. Надалі пропонується для цього поняття використовувати термін «метаслово», а також вводиться поняття «метафраза» стосовно фраз, які виникають внаслідок застосування метаслів. Загалом метаслова або метафрази описують різні варіанти власних назв, часу, напрямку, місця тощо.

Запропонована структура системи усного перекладу в межах предметних областей подана на рис. 1. За цією схемою спочатку користувач має обрати одну або декілька предметних областей. На сьогодні пропонується 15 ПО, з якими може працювати користувач. Активатор завантажує обрані ПО, що включає підсловник ПО з відповідними до цієї області типами речень (ТР), підсловник метаслів (метафраз), а також граматику, за якою моделюються допустимі обмеження на послідовності слів у реченнях.

Диктор вимовляє мовою 1 деяке речення, що розпізнається з урахуванням акустичної моделі (AM) та лінгвістичної моделі (LM). При цьому використовується акустичний компаратор, заснований на генеративній моделі, а AM і LM формуються відповідно до словників слів, метаслів та граматики обраних ПО.

Багатозначна відповідь розпізнавання у вигляді *n*-найкращих (*n*-Best) послідовностей слів порівнюється зі згенерованими допустимими моделями речень згідно з ТР. Далі приймається рішення про належність відповіді розпізнавання до певного ТР з урахуванням як акустичних мір схожості вхідного сигналу з еталонами *n*-найкращих послідовностей слів, так і міри відповідності послідовності слів деякому ТР. За знайденим найкращим ТР визначається тип змісту (ТЗ), а інтерпретатор знаходить відповідний ТЗ іншою мовою. На виході ми отримуємо текст мовою 2, який може бути озвученим відповідною системою перетворення текстів мови 2 на звук.

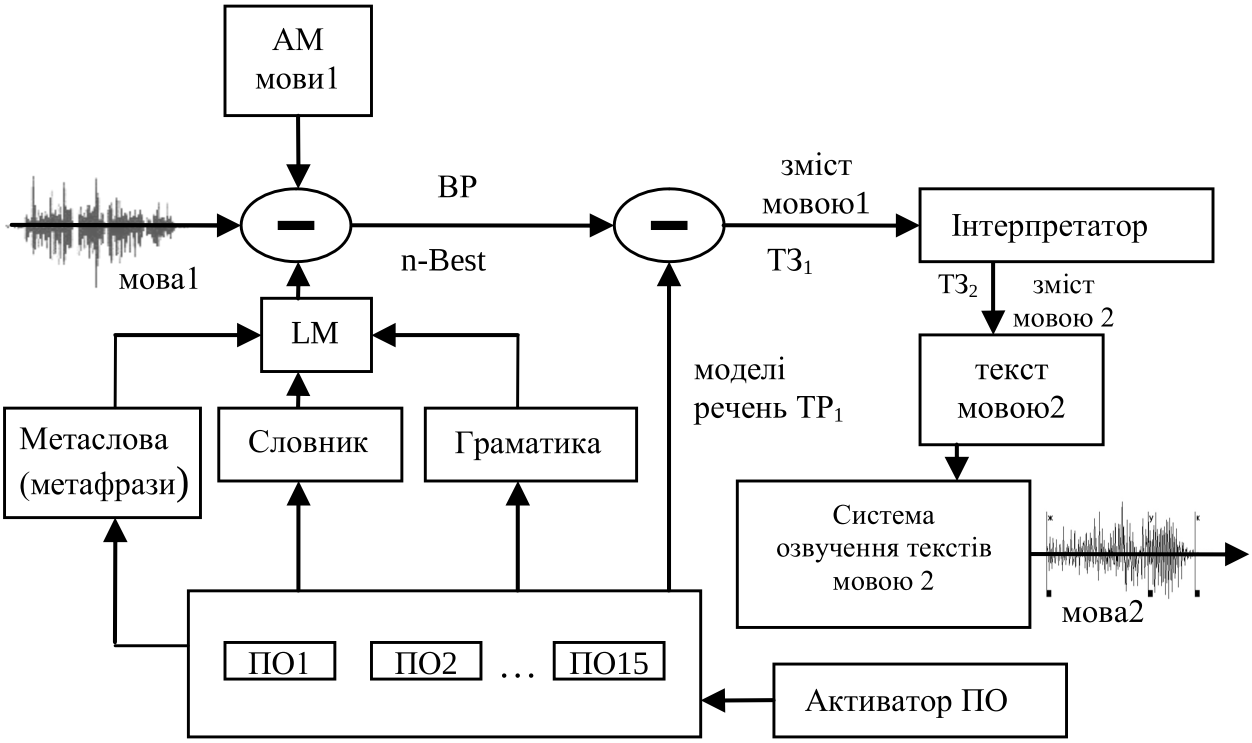


Рисунок 1 – Структура системи усного перекладу в межах предметних областей

В описаній структурі перекладу найменш досліджено процес переходу від відповіді розпізнавання до типу змісту. Шляхи розв’язку цієї задачі, описані в [3], ґрунтуються як на генеративній, так і на дискримінантній моделі.

Далі буде розглянуто загальну структуру засобів формування БД і З, що дає змогу реалізувати описану систему усного фразника-перекладача.

Структура БД і З усного фразника-перекладача

Представлена в роботі система усного перекладу є електронним аналогом паперового розмовника, взаємодія з яким відбувається найбільш природним способом – голосом. При розпізнаванні вимовленої користувачем фрази використовуються лінгвістичні та семантичні знання щодо обраної предметної області. Семантична модель, що описує предметну область і дозволяє відповідати на такі запитання з цієї предметної області, відповіді на які в явному вигляді не присутні в базі даних, являє собою базу знань. База знань є основним компонентом інтелектуальних та експертних систем.

Для системи усного перекладу (фразника-перекладача) дані, якими оперує користувач, мають бути зв’язаними, організованими за певними правилами, що передбачають загальні принципи опису, зберігання і маніпулювання, незалежно від прикладних програм. Саме такі дані розглядаємо (подаємо) як базу даних, формування якої виконується в межах предметних областей.

На рис. 2 подається функціональна схема, яка описує структуру засобів формування БД і З для фразника-перекладача. Ці засоби складають інструментарій, що дає змогу експерту оперувати предметними областями, незалежними та залежними від мови структурами, що відповідають за наповнення предметних областей.

Пропонована структура розрахована на реальну багатомовність. Перейшовши до деякої мови 1 (рідна мова, якою вимовляється фраза) та визначивши предметну область, експерт отримує доступ до типів смислів і речень, закладених у БД і З. Потім, обравши мову 2 (мова для перекладу фрази), експерт оперує варіантами перекладу вибраного типу речення.

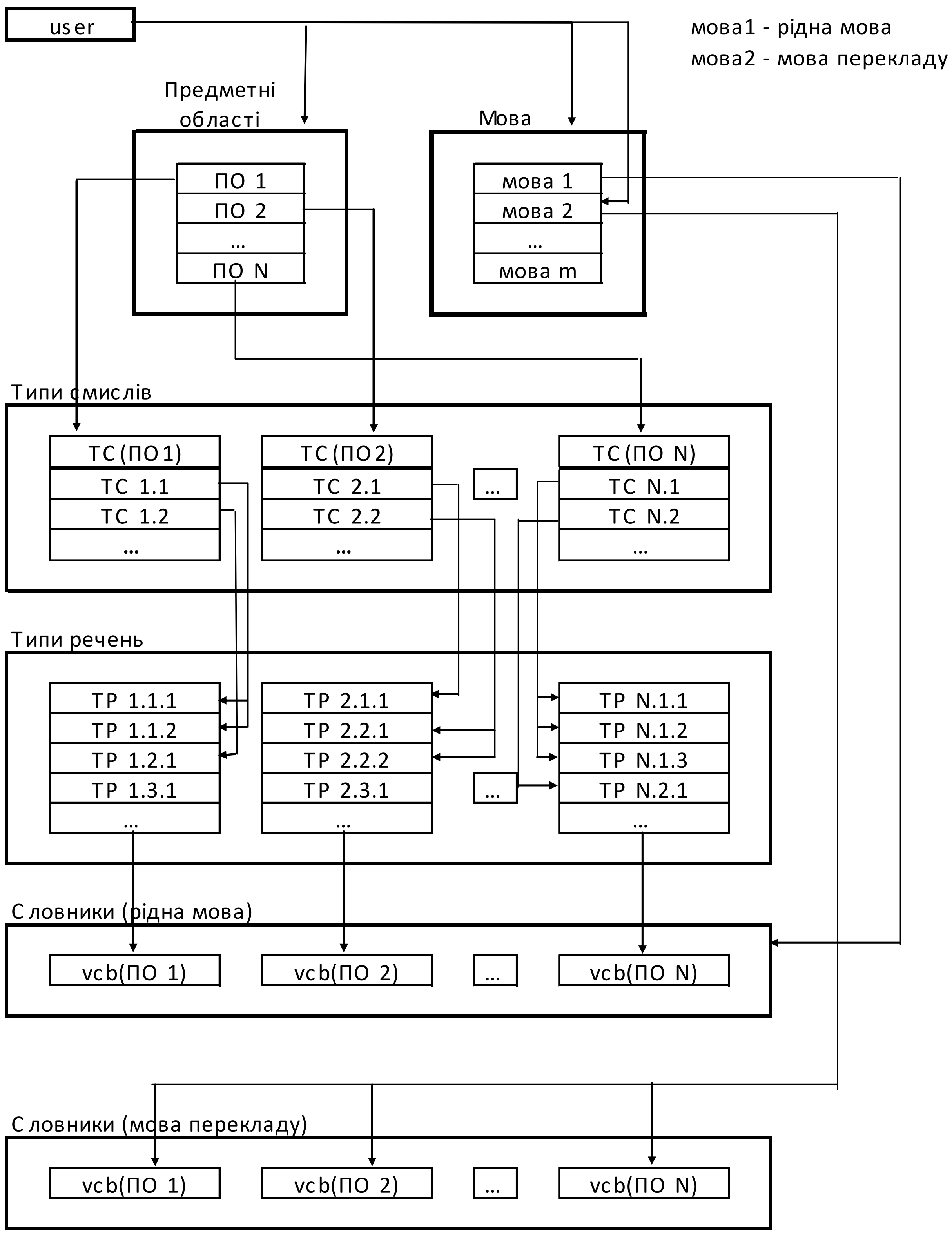


Рисунок 2 – Загальна структура БД і З усного фразника-перекладача

Новий тип речень заданої мови створюється шляхом додавання деякого типового зразка речення, що відповідає цьому ТР. При цьому цей зразок може містити метаслова або метафрази. Також передбачається виконання ряду інших дій: вилучення, редагування, переміщення в іншу ПО та збереження.

Автоматично формується список слів предметної області. Виходячи з цього списку, створюється словник вимови, який використовується акустичним компаратором у процесі розпізнавання для побудови композитних акустичних моделей слів з акустичних моделей фонем. При формуванні словника вимови застосовуються знання про закономірності зв’язку між фонетичними та орфографічними символами та позиціо-

вання наголосів у словах [5]. Параметри акустичних моделей фонем оцінюються засобами навчання/самонавчання генеративної моделі розпізнавання звукових образів [6], [7] на підставі мовленнєвого корпусу.

Окремо створюється словник метаслів і метафраз. Кожне метаслово є множиною слів. Метафраза складається з послідовностей, які містять як слова, так і метаслова. У цьому випадку забезпечується заборона циклічних вкладень.

Інший варіант використання БД і З – це пошук слова/метаслова у словнику для заданої мови та ПО і перегляд типів речень, де це слово зустрілося. Експерт обирає потрібне речення та аналізує переклад (мета)слова або речення іншою мовою.

У словниках вимови експерт-фонетист може переглядати та коригувати наголос слова та різні варіанти вимови (фонемні транскрипції).

Далі наводиться приклад реалізації системи автоматизованого формування БД і З для задачі усного перекладу в межах предметних областей у вигляді інструментарію. Експерт готує базові структури стандартними засобами редагування і далі використовує систему команд, яка здійснює необхідні перетворення.

Інструментарій формування бази даних і знань

Для формування БД та БЗ було розроблено інструментарій, який складається з комплексу модулів: *sentype2lattice, choossentences, sentencesTypes, trlitfile, text2phon, vcb2grammar, xlat_phones.*

На вході подається текстовий файл, який містить перелік типових фраз, що відповідають типам речень. Для модифікації, в якій передбачено генерування гіпотез речень, відповідними дужками експерт виділяє сегменти, що допускають перестановку. Цей файл повинен мати назву типу: thema1_struc_ua.txt. На виході отримуємо заповнені структури, що використовуються системою усного перекладу в межах предметних областей. Типове використання інструментарію передбачає декілька етапів.

1. Формування словника за текстовим файлом.

Спочатку запускається модуль *sentype2lattice*, за допомогою якого ми з представленої множини речень отримуємо всі можливі речення в мові діалогу, шляхом перестановок відповідних дужок. Запуск модуля виконується з командного рядка: sentype2lattice thema1_struc_ua.txt thema1_cyr.lat. Результат записується у текстовий файл thema1_cyr.lat.

Для отримання словника запускається наступний модуль *choossentences* з командного рядка: choossentences -t vcb_thema1_st.lat thema1_cyr.lat out.txt. В цьому модулі використано ключ, який дає можливість у словнику, який формується, до кожного слова приписати номери типів речень, де це слово зустрілося в межах ПО (перелік ТР іде через «;»). Отриманий словник буде записано у файл vcb_thema1_st.lat. Тут файл out.txt – це список пронумерованих усіх ТР.

2. Побудова граматики.

При розпізнаванні злитого мовлення ми вводимо деякі обмеження на порядок слідування слів. Найбільш ефективна граматика будується на основі лінгвістичного поняття про фонетичне слово [5]. Щоб формалізувати цю граматику, розглядаються лінгвістичні поняття енклітиків та проклітиків, які формуємо з певних частин мови.

Для отримання частин мови використовується модуль *text2phon*, запуск якого виконується з командного рядка: text2phon -transcribe 1 -rulespaph t2p_map.txt -save_lex_info tests/pos.txt tests/vcb_thema_st.lat.

Ключ -save_lex_info задає файл (tests/pos.txt), куди записується результат у вигляді списку слів з відповідними частинами мови. tests/vcb_thema1_st.lat – словник з типами речень, отриманий на попередньому етапі.

Відповідно до інформації у файлі pos.txt потрібно відшукати і позначити енклітики та проклітики (ecl та pcl) і зберегти новий файл під іменем pos_pe.txt

Далі для побудови самої граматики використовується модуль *vcb2grammar*, який запускається з командного рядка: vcb2grammar pos_pe.txt pos_pe.lat

На основі файла pos_pe.txt, отриманого на попередньому етапі, формується файл pos_pe.lat, куди записується результат. Цей файл доповнюється структурою граматики у текстовому вигляді: (pau <[pcl] [prep | pcl] ntl [ecl]> pau) і зберігається під іменем gram_pe.gram.

3. Формування словника вимови.

Щоб сформувати словник з фонемними транскрипціями слів, використовується модуль *all2MainRate*. В результаті отримаємо файл з фонемами vcb_thema1.txt.pht

На основі файлів словника та фонем формується новий файл словника, який складається з двох стовпчиків: слово і його фонемна транскрипція. Результат зберігається під іменем vcb_thema1_phon.txt.

У фонемній транскрипції є позначення наголошених голосних великими літерами та м'яких приголосних з апострофом. Виникає потреба привести транскрипції до вигляду, сумісного з алфавітом фонем акустичної моделі (A – a1, b' – b1). Для цього використовується модуль *xlat_phones*, який запускається з командного рядка: xlat_phones vcb_thema_phon.txt vcb_thema_phon.vcb phones0.txt -f 1. Тут результат записується у файл vcb_thema_phon.vcb.

Додатково до інструментарію включено модуль багатозначного перетворення мовленнєвого сигналу на текст, що реалізує генеративну модель розпізнавання звукових образів [8]. При цьому використовуються словник вимови (vcb_thema_phon.vcb), граматика (gram_pe.gram) та акустичні моделі фонем, які формуються засобами навчання/самонавчання [7] на підставі мовленнєвого корпусу. Результат розпізнавання буде записано у файл recout.mlf.

Для отримання перекладу використовується модуль *sentencesTypes*. Вхідними даними для цього модуля є vcb_thema1_st_lat – файл, що був отриманий на першому етапі, містить словник слів та список типів речень, де ці слова зустрічалися; thema1_en.txt – текстовий файл, що містить фрази англійською мовою, які відповідають українським фразам; thema1_struc_ua.txt – текстовий файл з українськими фразами, представленими у вигляді структур, та файл recout.mlf, який є результатом розпізнавання.

Запускається цей модуль з командного рядка: sentencesTypes -t vcb_thema_st_lat -ts thema_en.txt -ss thema_struc_ua.txt recout.mlf out_st.

В результаті буде отримано файл out_st, який буде містити фрази, продиктовані українською мовою, та їх відповідний переклад англійською мовою.

В наступному розділі описується розроблена система усного фразника-перекладача, аналізуються експериментальні дослідження використання цієї системи та наводяться результати її апробації.

Експериментальні дослідження

Для апробації та дослідження розробленої комп'ютерної моделі було створено дві системи: експериментальну та демонстраційну.

Експериментальна система складається з комплексу програмних модулів, описаних у попередньому розділі. В рамках цієї системи розпізнавання фраз (речень) проводилося на основі пофонемного декодера за умов обмеженої (на основі LISP-структур), вільної та відносно вільної (на основі фонетичних слів) послівних граматик [7]. Для оцінки параметрів акустичних моделей фонем використовується україномовний мовленнєвий корпус УкРеко, що містить записи 30 000 слів від більше 100 дикторів.

Цей корпус створений завдяки Гранту Президента України для обдарованої молоді № 58 за Розпорядженням Президента України від 14 грудня 2005 р. № 1276/2006-рп. Словники вимови сформовано для української та російської мов з апроксимацією вимови російських слів українськими фонемами.

Розроблена експериментальна система використовується для оцінки надійності, протоколів випробовування.

Цю систему реалізовано у вигляді модулів, що активуються командним рядком. Результатом роботи модуля є текст типу речення мовою, якою здійснюється усний переклад, тобто англомовний текст, перекладений з української мови. Цей текст надалі може бути озвучено шляхом відтворення стисненого запису або засобами синтезу мовлення за текстом.

Для експерименту довільним чином було вибрано 500 фраз, до яких було застосовано алгоритми пофонемного розпізнавання мовленнєвих сигналів в умовах обмеженої та відносно вільної (на основі фонетичних слів) послівних граматик [7]. При оцінюванні акустичних параметрів фонем використовувався мовленнєвий корпус кооперативу дикторів. Прийняття рішень щодо смислової інтерпретації здійснювалося згідно зі статистичним оцінюванням [4]. Результати експериментів наводяться в табл. 1.

Таблиця 1 – Результати розпізнавання та смислової інтерпретації 500 речень з двох предметних областей

Тип граматики	Надійність розпізнавання (%)		
	слів	речень	ТЗ
Обмежена	97	94,1	98,3
Вільна послівна	54,3	4,2	86,1
Відносно вільна	79,1	20,8	96,2

З наведених у табл. 1 результатів проведених експериментів можна зробити висновки, що для двох типів граматик (обмеженої і відносно вільної) похибка змістовної інтерпретації (ТЗ) не перевищує 4%, що є прийнятним для прикладної системи.

За умов обмеженої граматики швидкість розпізнавання у 7 разів перевищувала реальний час, а за умов вільної та відносно вільної граматик розпізнавання відбувалося швидше реального часу на ресурсах нетбука.

Для демонстрації роботи програмних модулів на основі описаної функціональної схеми та проведених досліджень було розроблено демонстраційну програмну модель з графічним інтерфейсом користувача (GUI) для перекладу речень в межах предметних областей (рис. 3).

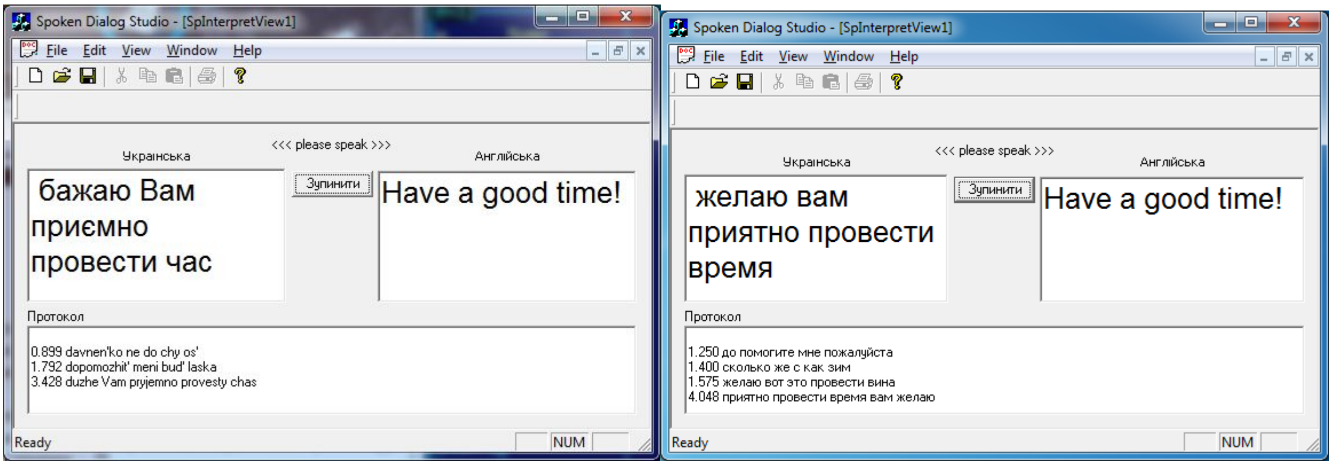


Рисунок 3 – Демонстраційна система усного фразника-перекладача з GUI

Запропонована демонстраційна система на даний час оперує трьома мовами (українська, російська та англійська), має 15 предметних областей, спільний словник для всіх ПО складає 2400 слів. Функціонує введення мовленнєвого сигналу для української та російської мов. При цьому слідування слів у вимовленій фразі може бути будь-яким з допустимих, включно з елементами спонтанності. Для введеного речення знаходиться найбільш подібний до нього тип речення та відповідний тип змісту, а перше речення цього типу змісту оголошується результатом перекладу.

Створена система усного фразника-перекладача з графічним інтерфейсом, що дає змогу перекладати на англійську мови з української або з російської, успішно демонструвалася на ряді заходів. Це відкриває широкі можливості застосування розробленої системи для нових мов.

Висновки

Система усного перекладу, яка представлена в роботі, побудована за прикладом паперового розмовника. Дана система передбачає взаємодію з електронним пристроєм за допомогою голосу. При розпізнаванні вимовленої користувачем фрази використовуються лінгвістичні та семантичні знання щодо обраної предметної області. Введення обмежень при формуванні граматик та використання метаслів розширюють можливості системи та підвищують надійність розпізнавання. Таким чином, користувач отримує можливість більш вільно та розмаїто спілкуватися в іншомовному середовищі.

Розроблений інструментарій є ефективним засобом розширення предметних областей і додавання нових мов для системи усного перекладу. Експерту-лінгвісту, який наповнює БД і З, достатньо володіти мінімальними навиками роботи на комп'ютері. Заплановане введення графічного інтерфейсу користувача має покращити взаємодію експерта з інструментарієм, що загалом підвищить продуктивність праці.

Актуальним залишається питання інтонації та розділових знаків. Одні і ті ж самі фрази, вимовлені з різною інтонацією, можуть виражати як питальне речення, так і розповідне. Тому в подальшій роботі слід дослідити можливість розпізнавання інтонації та ритму (просодики) з метою автоматичного розставляння розділових знаків у розпізнаних фразах.

Планується ввести відмову від інтерпретації фрази у разі невідповідності жодному типу речень та формувати більш точний відповідник мовою перекладу, аналізуючи допустимі типи речень та враховуючи метаслова. Аналіз контексту на рівні типу змісту дасть змогу моделювати багатомовний усний діалог.

Література

1. Режим доступу : translate.google.com.ua
2. Сажок Микола. Система усного перекладу на основі інтерпретації мовленнєвого сигналу в межах предметних областей / Микола Сажок, Валентина Яценко // Всеукраїнська міжнародна конференція з оброблення сигналів і зображень та розпізнавання образів УкрОбраз'2010. – Київ, 2010. – С. 103-106.
3. Яценко В.В. Параметризація типів речень предметної області для системи усного фразника-перекладача / В.В. Яценко // Штучний інтелект. – 2011. – № 4. – С. 134-142.
4. Sazhok M. Comparison of word grammars for spoken translation in subject area / M. Sazhok, V. Yatsenko // Proceedings of the 13th International Conference on Speech and Computer – SpeCom'2009, (St. Petersburg, Russia, 2009). – P. 494-497.
5. Робейко В.В. Багатозначна багаторівнева модель перетворення орфографічного тексту на фонемний / В.В. Робейко, М.М. Сажок // Штучний інтелект. – 2011. – № 4. – С. 117-125.
6. Винцюк Т.К. Анализ, распознавание и смысловая интерпретация речевых сигналов. – Киев : Наукова думка, 1987.

7. HTK Book, version 3.1 / Young S.J. et al. – Cambridge University, 2002.
8. Lee A. Recent Development of Open-Source Speech Recognition Engine Julius / A. Lee, T. Kawahara // Asia-Pacific Signal and Information Processing Association Annual Summit and Conference (APSIPA ASC). – 2009.

Literatura

1. Sazhok M. UkrObraz'2010. Kyiv, 2010. S. 103-106.
2. Yatsenko V.V. Shtuchnyj intelekt. 2011. № 4. S. 134-142.
3. Vintsiuk T.K. Analiz, raspoznavaniye i smyslovaya interpretatsiya rechevykh signalov. Kiev: Naukova Dumka. 1987.
4. Young S.J. et al., HTK Book, version 3.1, Cambridge University. 2002.
5. translate.google.com.ua
6. Robeiko V.V. Shtuchnyy intelekt. Donets'k. 2011. № 4. S. 117-125.
7. Sazhok M.. Comparison of word grammars for spoken translation in subject area. SpeCom'2009. S. 494-497.
8. Lee A. KawaharaT.. "Recent Development of Open-Source Speech Recognition Engine Julius" APSIPA ASC. 2009.

RESUME

V.V. Yatsenko, M.M. Sazhok

Automated Means for Formation of Linguistic Data and Knowledge Base for Spoken Translation System

This paper describes the formation of data and knowledge base for spoken interpreter system. The spoken interpreter system is based on the generative model for speech recognition and restricted to specific subject areas. Input speech signal is being converted to sequences of words for the source language. Then the decision about the respective meaning type is taken proceeding from comparison of recognized word sequences to sentence models called sentence types. Therefore, an incorrect recognition response does not mean an incorrect understanding response. To model variety of proper names, time, place, etc. we exploit meta-words and meta-sentences.

The proposed structure of data and knowledge base allows for operating language dependent and language independent structures. A subject area represents the semantic level and consists of meaning types which are the explicit understanding units. Each meaning type contains sentence types modeling all permissible sentences that express same meaning. The correspondence between sentence types of available languages is provided. To implement the structure creation and update a toolkit is developed and described. Thus, an expert-linguist may use the straightforward toolkit to form a linguistic data and knowledge base for new subject areas and languages.

The experimental system of a spoken phrasebook-interpreter exposed results of less than 4% meaning error rate that is proper for practical applications. The presented GUI-based demo-version is provided with the data and knowledge base for a new source language created by the developed means. Multilingual spoken dialog system modeling is the next stage of the research.

Стаття надійшла до редакції 05.07.2012.